Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-348

Back analysis of the Old Oak Common box using surrogate models

Yunxiang Yang, David M G Taborda, Aikaterini Tsiampousi
Imperial College London, London, UK, yunxiang.yang22@imperial.ac.uk

Agustín Ruiz López
Seequent – The Bentley Subsurface Company, Delft, The Netherlands

Antonio M G Pedro
Universidade de Coimbra, ISISE, ARISE, Coimbra, Portugal

Stuart Hardy
ARUP, London, UK

ABSTRACT: The observational method (OM) has been adopted in urban excavation projects to help manage uncertainties associated with design soil parameters. By deploying comprehensive ground and structural monitoring schemes, the OM can provide a continuous assessment of the progress of the project, with deviations between observed and predicted responses potentially leading to an interruption of construction operations and/or a review of the original design. As part of this process, the OM requires the frequent back analysis of soil parameters using field monitoring data, with the obtained updated soil parameters being subsequently used to simulate future construction stages. This enables the implementation of remedial measures to be evaluated, as well as potential cost saving options to be explored. In this context, the numerical model must reliably capture the actual soil behaviour while remaining computationally efficient, as the back analysis process typically requires many simulations to identify optimal parameter sets. However, such high computational costs are generally incompatible with the constraints imposed by construction timelines and cost considerations, thereby restricting the feasibility of employing high-fidelity numerical models. To address these challenges, this paper utilises an Artificial Neural Network (ANN) as a surrogate model, capable of predicting results at different construction stages comparable to those generated by high-fidelity numerical simulations with significantly reduced computational effort, in conjunction with Genetic Algorithms (GA) for automatic back analysis of soil parameters based on field data. The proposed approach for automatic back-analysis is demonstrated through a retrospective case study of the Old Oak Common station in west London (United Kingdom) which is part of the High Speed 2 (HS2) project, highlighting its ability to identify model parameters based on field measurements.

KEYWORDS: Machine learning, surrogate modelling, observational method, back analysis, urban excavations.

1 INTRODUCTION

The observational method (OM), originally introduced by Karl Terzaghi (Peck, 1969), was developed to address the considerable uncertainties associated with soil parameters during design in Geotechnical Engineering. These uncertainties primarily arise from the variability in natural materials, limitations of site investigation and laboratory testing, and assumptions inherent to engineering analysis. By comparing field monitoring data systematically collected during construction with design predictions, the OM supports engineers in making informed modifications to construction operations. These can include remedial measures to ensure the safety of the project or, should the observed response suggest that the soil-structure system is more competent than anticipated, changes to the construction sequence. The latter may, for instance, involve the removal of propping levels (Yeow et al., 2014). This adaptive approach not only enhances project reliability but also offers considerable time and cost savings (Hardy et al., 2018; Powderham & O'Brien, 2020).

Within the framework of the OM, soil parameters are updated through back analysis as soon as new field monitoring data become available. While optimisation techniques, such as Genetic Algorithms, facilitate the automation of this procedure, many numerical simulations are required, potentially making it prohibitive under the time constraints of real-world construction projects. Surrogate models can generate results comparable to those of high-fidelity numerical models, such as those based on the Finite Element Method (FEM), at a fraction of the computational cost. This paper presents the back analysis of an excavation in London Clay by integrating surrogate models with optimisation algorithms, thereby enabling efficient parameter identification.

2 CASE STUDY

2.1 *Description of the Old Oak Common station*

Old Oak Common station is a new railway station constructed as part of the High Speed 2 (HS2) project in London. The station is approximately 840 m long, 16 to 66 m wide, and up to 20 m deep, constructed using a typical top-down approach. This paper focuses on design Section 10, which features a 1200 mm thick diaphragm wall extending to a depth of 40.5 meters, supported during excavation by slabs at depths of 3.55 m ('ground level') and 9.75 m ('intermediate level'). The application of the OM in this project was to assess whether a third temporary support prop at a depth of 16.5 m, i.e. between intermediate and formation level, could be omitted ('temporary level'), which required extensive back analysis of the movements observed prior to this stage. The field data include measurements of the diaphragm wall's horizontal deformation relative to its toe at key excavation stages: intermediate prop level, temporary prop level, and formation level, with other site observations not available for use in this analysis.

The ground surface at this site is located at 26.5 mOD and the stratigraphy consists of 1 m of Piling Material, followed by 5 m of Weathered London Clay, 65 m of London Clay, 17.5 m of Lambeth Group Clay, 2.5 m of Lambeth Group Sand and a 2 m thick layer of Thanet Sand, extending down to -67.5 mOD. Below this level, as typically found in London, is chalk. The unit weight of the Piling Material is 19 kN/m^3, whereas all the other soils have an assumed unit weight of 20 kN/m^3.

2.2 *Reference numerical analysis*

The simulation was performed using the finite element software PLAXIS 2D 2024.02 (Bentley Systems, 2024), under planestrain conditions with the lateral boundary at the far field positioned 100 m from the diaphragm wall and the base of the model coinciding with that of the Thanet Sand. The pore water pressure (pwp) profile on site showed signs of underdrainage which was imposed as an initial condition to the problem by performing a steady state analysis, with the perched water table located at the base of the Piling Material and the lower aquifer at a depth of about 51.5 m. The permeability of the London Clay and Lambeth Group Clay was modelled as both anisotropic ($k_h = 2 \cdot k_v$, where k_v and k_h are the vertical and horizontal permeabilities, respectively) and varying with depth. This was implemented as a User-Defined Flow Model in PLAXIS (Taborda et al., 2023) and calibrated using data from Hight et al. (2007). The logarithmic variation of k_v with depth was described by the following expression:

$$log_{10}k_v = 0.0215 \cdot (y - 25.5) - 9.427 \quad (1)$$

where y is the vertical coordinate in mOD. The simulated profile, illustrated in Figure 1, showed good agreement with that measured in situ. Also shown in Figure 1 is the adopted profile of coefficient of earth pressure at-rest (K_0).

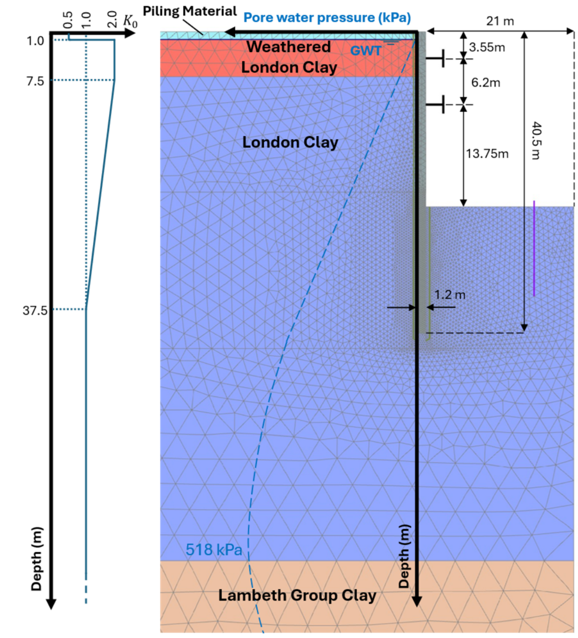

Figure 1. Geometry and ground conditions for OOC Section 10.

Table 1. Adopted strength (φ, c) and dilatancy (ψ) properties.

Material	$\varphi(°)$	$c\ (kPa)$	$\psi\ (°)$
Piling Material (PM)	25.0	3.0	0.0
Weathered London Clay (WLC)	23.0	0.0	12.5
London Clay (LC)	25.0	5.0	12.5
Lambeth Group Clay (LGC)	28.0	5.0	13.5
Lambeth Group Sand (LGS)	34.0	0.0	17.0
Thanet Sand (TS)	37.0	0.0	20.0

All soils were modelled with a non-associated Mohr-Coulomb failure criterion, with strength and dilatancy parameters listed in Table 1. Non-linear elastic stiffness was simulated for all materials using the IC MAGE M01 model (Taborda et al., 2024), except for the Piling Material, which was modelled with a linear elastic stiffness ($E = 10\ MPa,\ \nu = 0.2$). The stiffness parameters for M01 were adopted from Sailer et al. (2019), based on the back-analysis of a retaining wall in London. For simplicity, it was assumed that Weathered London Clay and London Clay had the same stiffness properties. The diaphragm wall was modelled as wished-in-place using solid elements to which concrete properties were assigned from the design report ($E = 30\ GPa,\ \nu = 0.2$). The props were modelled with EA values of $6\ GPa$ and $4.4\ GPa$ for the 'ground' and 'intermediate' levels, respectively. The soil-structure interaction was simulated using interface elements, with interface strength reduced to two-thirds of that of the surrounding soil strength and stiffness governed by a power law assuming a reduction factor of 1.0, with parameters obtained from the modelled G_{max} profile, following the procedure outlined in Bentley Systems (2024).

The excavation was assumed to take place under undrained conditions for the clay layers, reflecting their relatively low soil permeability in comparison to the construction timeframe. The excavation was performed in steps of 2 to 3 metres per layer.

2.3 *Back-analysis*

The back-analysis focused exclusively on the stiffness parameters of London Clay, as the excavation takes place fully within this material, assuming, as before, that London Clay and Weathered London Clay layers shared identical stiffness properties. The four properties varied were G_{ref} (stiffness at very small strains), a (strain level at which stiffness reduction takes place), b (nonlinearity rate of strain reduction) and $R_{G,min}$ (ratio G_{min}/G_{max} at very large strains). Sufficiently broad intervals were adopted for the back-analysis to encompass the range of soil conditions reported in London (Hight et al., 2007; Jurečič, Zdravković & Jovičić, 2013), as detailed in Table 2. The interface element stiffness was adjusted in accordance with the varying LC parameters for each numerical analysis.

Table 2. Back analysis intervals for London Clay.

Bound	$G_{ref}\ (kPa)$	a	b	$R_{G,min}$
Lower	16667	1E-5	0.5	0.01
Upper	61667	1E-2	1.5	0.2

3 SURROGATE MODEL

This study used a single Artificial Neural Network (ANN) in *Keras* (Chollet et al., 2015) to predict diaphragm wall deformation at both intermediate (9.75 m) and temporary (16.5 m) prop levels during excavation. Training one ANN for multiple stages was expected to enhance back analysis and overall results compared to separate models. The adaptive sampling method CV-BASHES was applied, which dynamically constructs the training set by balancing global exploration with local exploitation. Cross-validation error guides new sample generations toward regions of high model uncertainty, and candidate points are ranked by predicted error so that high-error regions are refined iteratively to improve surrogate accuracy.

3.1 *Synthetic database*

Two synthetic databases, comprising 40 samples for initial training and 50 samples for consistent testing, were generated for CV-BASHES using the Latin Hypercube Sampling method (Virtanen et al., 2020). Each sample (i.e. a full finite element simulation) corresponds to a model as described in Section 2.2, with distinct combinations of G_{ref}, a, b and $R_{G,min}$ from within the bounds listed in Table 2.

Prior to the training of the ANN surrogate, the diaphragm wall's deformation profile was discretised into 40 equally spaced points along its height. This approach resulted in each sample containing 80 training points (i.e. 40 for each of the two excavation stages), with identical values of G_{ref}, a, b and $R_{G,min}$, but different elevation y/H, horizontal deformation (u_x)

and phase indicator, the latter distinguishing between results for 'intermediate' and 'temporary' levels. To ensure effective training and improve the generalisation ability of the ANN model, normalisation was applied to the dataset so that all variables were on a comparable scale.

3.2 *ANN training*

The feedforward ANN model has six input features, with the horizontal deformation being its single output. During each CV-BASHES iteration, 10 new samples were created and the ANN's hyperparameters were retuned within the ranges specified in Table 3. The *Adam* optimiser and early stopping were applied to enhance the robustness of the model. For validation, data splitting was conducted such that all training points from the same sample were assigned to the same fold. CV-BASHES stopped generating new training samples once the ANN achieved sufficient accuracy.

Table 3. Hyperparameter tuning ranges for ANN.

Hyperparameter	Range
Number of layers	[3,6]
Number of neurons per layer	[500, 1500]
Activation function	{'relu', 'leaky_relu', 'elu'}
Batch size	[16, 128]
Learning rate	[10^{-6}, 10^{-2}]

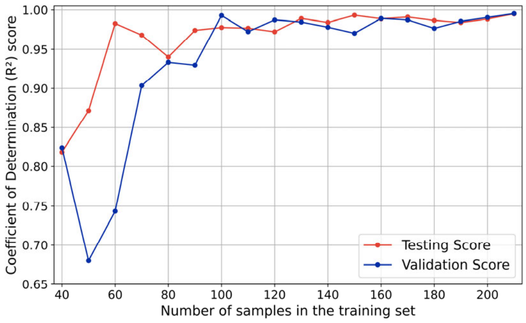

Figure 2. Learning curve of the ANN surrogate with CV-BASHES.

3.3 *ANN performance evaluation*

The final architecture of the ANN surrogate comprised three layers, each containing 624 neurons with the Exponential Linear Unit (ELU) activation function. The model was trained with a learning rate of 3.776×10^{-5} and a batch size of 40. Figure 2 illustrates the learning curve of the ANN surrogate, showcasing progressive performance enhancement after each new CV-BASHES iteration. The adaptive sampling procedure converged after 18 iterations, yielding a final training database of 210 samples. In the first two iterations, testing accuracy increased, while validation scores showed a temporary decline. This behaviour exemplifies the dynamics of the adaptive sampling process: newly adapted samples enhanced the overall model performance when incorporated into the training set, yet they simultaneously posed a predictive challenge as validation samples. In subsequent CV-BASHES iterations, both testing and validation metrics improved. In the final iterations, there was a close agreement between the testing and cross-validation scores, demonstrating the robustness of the final ANN, thereby confirming the sufficiency of the database size and justifying the termination of the adaptive sampling procedure.

The final ANN achieved a coefficient of determination (R^2) of 0.998 for training and 0.995 for testing dataset. Figure 3 compares the ANN predictions for the testing dataset (black markers) with the "ideal fit" line (1:1, red dashed). Out of the 50 testing samples, 49 exhibited excellent agreement with the ideal 1:1 fit, while only one sample showing discrepancies, though limited to approximately 10%.

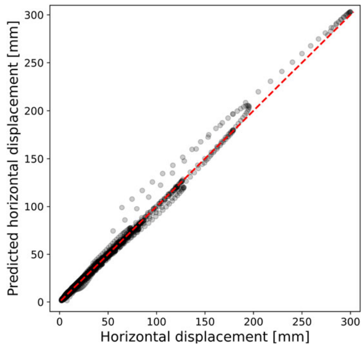

Figure 3. Regression plot for the testing dataset.

4 BACK-ANALYSIS

Back-analysis is fundamentally an optimisation procedure aimed at identifying the global optimal set of soil parameters that provide the closest match with a given observed field response. In this study, the objective is to calibrate the M01 stiffness parameters for WLC and LC layers, using the OOC Section 10 field monitoring data during excavation to 'ground' and 'intermediate' levels. Once calibrated, the back-analysed soil parameters are used to predict the horizontal deformation of the diaphragm wall at formation level. Through comparison with the monitoring data for the final phase, the validity of the obtained stiffness parameters can be further assessed.

4.1 *Genetic Algorithms with ANN surrogate*

Genetic Algorithms (GA) have shown promising performance when applied to back-analysis problems in geotechnical engineering (Ferrero et al., 2023). In this study, GA was executed five times, with a search space identical to the ANN sampling domain (Table 2). The adopted GA parameters are summarised in Table 4. The value of R^2 was employed as the fitness function, quantifying the agreement between the measured horizontal deformation profiles and the predictions produced by the ANN surrogate model.

Table 4. Parameters adopted in the Genetic Algorithm.

Parameter	Value
Population size	1000
Number of generations	200
Mutation probability	0.3
k-Tournament (k value)	3
Elitism	5
Maximum generations allowed without fitness improvement	20

4.2 *Results and forward prediction*

Table 5 presents the three best results from multiple GA realisations. While optimal parameter sets vary considerably, each achieves a good agreement with field observations. This ambiguity suggests that the four parameters of the adopted nonlinear stiffness model, despite having clear definitions in terms of their impact on the obtained stiffness reduction curve, introduce excessive degrees of freedom in comparison with the available constraints (the two measured displacement profiles). However, as can be seen in Figure 4, the three sets of parameters result in very similar stiffness values for strains between 0.02% and 0.2%. This range is similar to that suggested by Atkinson & Sallfors (1991) for stiffness mobilisation in stiff clays supporting retaining walls.

Table 5. Best solutions returned by the Genetic Algorithm.

Rank	G_{ref} (kPa)	a	b	$R_{G,min}$	R^2
1	56981	0.000310	1.499	0.0430	0.9577
2	53375	0.000365	1.492	0.0396	0.9567
3	46161	0.000565	1.497	0.0199	0.9550

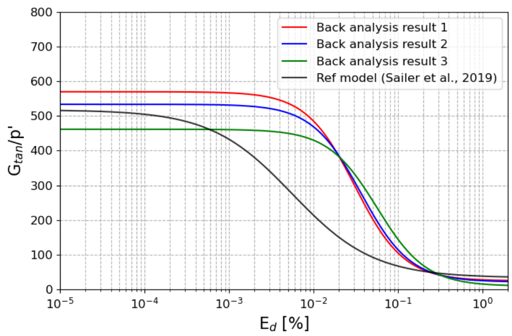

Figure 4. Comparison between stiffness reduction curves.

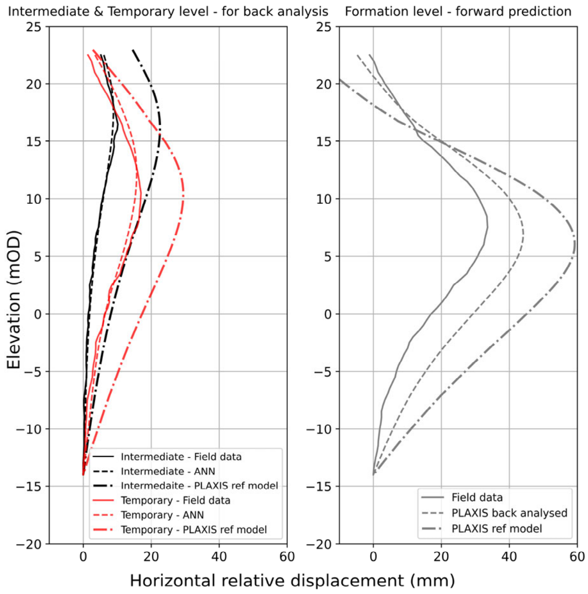

Figure 5. Comparison of the best back analysis result and the reference model versus field observations.

Figure 5 compares the results of best-fit parameter set, together with the reference model using stiffness parameters proposed in Sailer et al. (2019) and the monitoring data. Clearly, the new set of parameters can replicate very accurately the observed displacements. It is important to note that a single set of parameters is able to capture the response at both stages, demonstrating the advantages of using a model that can simulate the evolving soil stiffness with ground deformation. As expected from the comparison between back-analysed stiffness and original stiffness in Figure 4, the new parameters yield substantially lower movements than the reference analysis. In terms of the forward prediction for formation level, the new parameters lead to ground movements that are closer to the observed data (i.e. with only two prop levels), when compared with the results corresponding to the reference parameter set. However, there is still a slight overprediction, indicating that the stiffness at larger deformation levels should have been further reduced. This difference does not indicate that typical London Clay parameters are unsuitable but reflects the limitations of generic parameter sets for site-specific behaviour. These findings highlight both the effectiveness and limitations of the proposed approach.

5 CONCLUSIONS

In this study, London Clay stiffness parameters were back-analysed from excavation movements measured at the Old Oak Common station in London. The back analysis was carried out by coupling surrogate models with Genetic Algorithms (GA). An Artificial Neural Network (ANN) was trained on 210 samples using the adaptive sampling method CV-BASHES, resulting in a surrogate model capable of reproducing the results of high-fidelity FE analyses. The use of CV-BASHES meant that the number of samples did not need to be pre-determined, streamlining the creation of the surrogate model. By integrating the ANN with GA, back-analysis was conducted based on field observations from two key construction phases in a fraction of the time that would have been required if the full FE model had been used. The forward prediction derived from the back analysed soil parameters showed reasonable agreement with field measurements. Overall, the results demonstrate the potential of the adopted methodology. A key limitation is reliance on synthetic data; future work will incorporate field data and explore transfer learning for real-world applicability.

6 ACKNOWLEDGEMENTS

The work was supported by the Department of Civil and Environmental Engineering, Imperial College London, through a Dixon scholarship, and by EPSRC project SIDeTools [EP/X52556X/1].

7 REFERENCES

Atkinson, J. H. & Sallfors, G. (1991). Experimental determination of soil properties. General Report to Session 1. *Proceedings of the 10th ECSMFE*, Florence 3, 915-956.

Bentley Systems 2024. *PLAXIS 2D 2024.2 Reference Manual*. Bentley Systems, Exton, PA, USA.

Chollet, F., & others. 2015. Keras. https://keras.io/.

Ferrero, J. A., Ruiz Lopez, A., Taborda, D. M. G., Brasile, S. 2023. Applying the observational method to a deep braced excavation using an artificial neural network. *Proc. NUMGE 2023*, London, UK. http://doi.org/10.53243/NUMGE2023-303.

Jurečič, N., Zdravković, L. and Jovičić, V. 2013. Predicting ground movements in London Clay. *Proceedings of the Institution of Civil Engineers – Geotechnical Engineering*, 166 (5), 466-482.

Hardy, S., Nicholson, D., Ingram, P., Gaba, A., Chen, Y. and Biscontin, G. 2018. New observational method framework for embedded walls. *Geotechnical Research*, 5 (3), 122-129.

Hight, D.W., Gasparre, A., Nishimura, S., Minh, N.A., Jardine, R.J. and Coop, M.R. 2007. Characteristics of the London Clay from the Terminal 5 site at Heathrow Airport. *Géotechnique*, 57 (1), 3-18.

Peck, R.B. 1969. Advantages and limitations of the observational method in applied soil mechanics. *Geotechnique*, 19 (2), 171-187.

Powderham, A. & O'Brien, A. 2020. *The Observational Method in Civil Engineering: Minimising Risk, Maximising Economy.* CRC Press.

Sailer, E., Taborda, D.M., Zdravković, L. and Potts, D.M. 2019. Fundamentals of the coupled thermo-hydro-mechanical behaviour of thermo-active retaining walls. *Computers and Geotechnics*, 109, 189-203.

Taborda, D. M. G., Kontoe, S. and Tsiampousi, A. 2023. IC MAGE Flow Model 01 – Nonlinearly varying anisotropic permeability (Version 1.0). *Zenodo*. https://doi.org/10.5281/zenodo.7796383

Taborda, D. M. G., Kontoe, S. and Tsiampousi, A. 2024. IC MAGE Model 01 – strain-hardening/softening Mohr-Coulomb failure criterion with isotropic small strain stiffness (Version 2.2). *Zenodo*. https://doi.org/10.5281/zenodo.13891265

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & Van Mulbregt, P. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods*, 17 (3), 261-272

Yeow, H.-C., Nicholson, D., Man, C.-L., Ringer, A., Glass, P. and Black, M. 2014. Application of observational method at Crossrail Tottenham Court Road station, UK. *Proceedings of the Institution of Civil Engineers – Geotechnical Engineering*, 167 (2), 182-193.